Chapter 14
Tonal Languages and Cochlear Implants

Li Xu and Ning Zhou

1 Introduction

As a major part of world languages, tonal languages are spoken in every continent except for Australia. In a tonal language, voice pitch variation (i.e., tone) at the monosyllabic level is a segmental structure that conveys lexical meaning of a word (Duanmu 2000). Mandarin Chinese, a tonal language, is spoken by more people than any other single language, including non-tonal languages. While some dialects in southern Mexico may distinguish as many as 14 tones, Chinese dialects typically have 4–6 contrastive tones.

Multi-channel cochlear implants (CIs) have been a great success in providing profoundly deafened individuals with satisfactory speech perception in quiet. The contemporary speech-processing strategies deliver primarily temporal-envelope information of speech to the auditory nerve of the implantees (see Loizou 2006 for a review). These strategies do not explicitly code pitch information, because they have been mainly designed to accommodate Western languages that use pitch variation only for suprasegmental structures, such as intonation difference between a statement and a question. Because of the lack of pitch coding, tonal-language understanding remains challenging for implant users. The challenge can be tone recognition as well as tone production in language development. Both temporal and spectral approaches have been taken to improve CI pitch perception.

This chapter will describe the acoustical cues for recognition of lexical tones, primarily the Mandarin Chinese tones, and the relative contributions of these cues. This chapter will also discuss results in tone recognition in implant recipients in relation to their differences in demographics, devices, strategies, and psychophysics.

L. Xu (✉)
School of Rehabilitation and Communication Sciences,
Ohio University, Athens, OH 45701, USA
e-mail: xul@ohio.edu

F.-G. Zeng et al. (eds.), *Auditory Prostheses: New Horizons*, Springer Handbook of Auditory Research 39, DOI 10.1007/978-1-4419-9434-9_14,

The effects of frequency-place mismatch on tone recognition, which is a unique problem in CI users, will also be discussed. This chapter will then explore the relationship between music pitch perception and lexical tone recognition. Lastly, this chapter will evaluate results on tone production and vocal singing in prelingually deafened, native tonal-language speaking children with CIs.

2 Acoustical Cues for Tone Recognition

Mandarin Chinese has 4 lexical tones that are commonly known as tone 1, tone 2, tone 3, and tone 4. The Mandarin tones are classified based on both the fundamental frequency (F0) variation patterns and the absolute frequency heights (Howie 1976). Tone 1 has a "high flat" F0 pattern, and tone 2 has a "middle low rising" pattern. Tone 3 has a "dipping and rising" contour, with a possible break on the turning point from dipping to rising. The break reflects the loss of voicing that corresponds to a glottal stop. Sometimes tone 3 can also lose its final rising, resulting in a F0 contour that falls from a moderate level to a low level without rising. Tone 4 has a "high falling" contour. Figure 14.1 shows the F0 contours of the 4 Mandarin tones spoken by multiple talkers.

2.1 *Primary and Secondary Cues*

The F0 height and contours are the primary intrinsic cues for tone recognition. Speech materials contain redundant F0 information. Liang (1963) demonstrated that high-level tone recognition was preserved even when the Mandarin Chinese speech signals

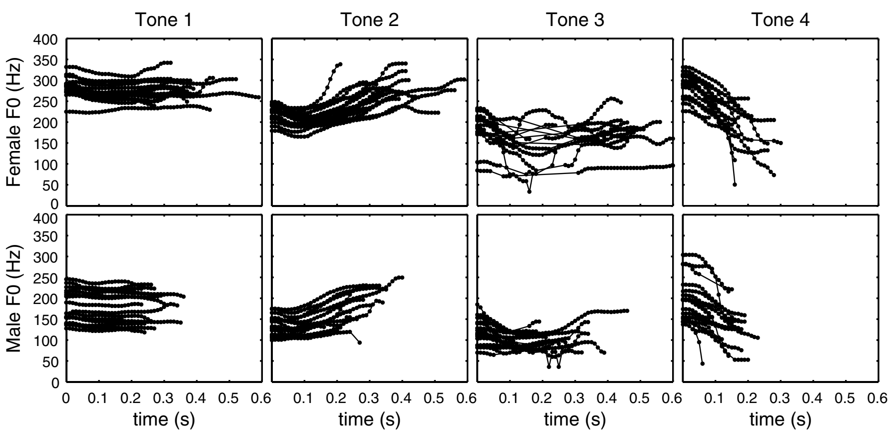

Fig. 14.1 Fundamental frequency (*F0*) contours of the four Mandarin tones produced by 16 female (*upper panel*) and 16 male (*lower panel*) native speakers of Mandarin Chinese. Adapted from Lee and Hung (2008) with permission from the Acoustical Society of America

were highpass filtered at 2.4 kHz, via presumably the unresolved harmonics inducing the F0 residue pitch (Schouten et al. 1962). The F0 contour itself is considered redundant for tone recognition; that is, not all parts of the contour are necessary for tone recognition. Liu and Samuel (2004) found that perception of tone 3 of Mandarin Chinese was not affected when the rising part of the F0 contour was neutralized. Gottfried and Suiter (1997) and Lee (2009) also showed the redundancy of F0 contour by producing better than chance performance with shortened tone stimuli including only the preceding consonant and the following 6 glottal periods of the vowel.

F0 constitutes the most important acoustic characteristic of tones, but secondary acoustic cues are also useful for tone recognition, particularly when F0 is compromised. These secondary cues include duration, amplitude contour, and spectral envelope of the speech signal.

Mandarin tones differ in duration, with tone 3 having the longest duration. Duration differences of the other 3 tones, however, are less consistent (e.g., Howie 1976; Luo and Wang 1981). Fu and Zeng (2000) recorded tones from 6 syllables spoken by 10 talkers and found that the average duration for tone 3 was the longest (463.3 ms), followed by tone 2 (374.7 ms) and tone 1 (339.5 ms), with tone 4 being the shortest (334.4 ms). Xu et al. (2002) and Lee and Hung (2008) found a similar distribution of tone durations (see Fig. 14.1).

The reliability of the duration cue for tone recognition is still in debate. Based on a maximum likelihood model (Green and Swets 1966), Xu et al. (2002) found that tone recognition can be as high as 56.5% correct with duration cues alone. Xu et al. (2002) reported that tone recognition reduced by approximately 15 percentage points for all 4 tones of equal durations as compared to that using the tokens with preserved durations in a vocoder study (more details below).

The other secondary cue is the overall amplitude contour or temporal envelope. It has been shown that tone recognition, using the signal-correlated noise with controlled duration, still remained above chance (Whalen and Xu 1992). Whalen and Xu attributed the performance to the energy distribution in the amplitude contour of the tones. Fu and Zeng (2000) further used signal-correlated noises to explore the roles of amplitude contour, duration, and periodicity in tone recognition. Tone recognition was around 70% correct, when all 3 temporal cues were available. Either the amplitude contour or the periodicity cue alone resulted in approximately 55% correct recognition, and the duration cue alone provided the lowest recognition score of about 35% correct. These results were confirmed in a more recent acoustical and perceptual study by Kuo et al. (2008).

A third secondary cue is related to the speech spectral envelope. In whispered speech, where the spectral envelope is preserved but the voice source is absent, Mandarin Chinese tone recognition is fairly good (i.e., 60–70% correct) (Liang 1963). Kong and Zeng (2006) pointed out that the duration and overall amplitude cues alone cannot account for the significant tone-recognition performance of whispered speech. They reasoned that the formant frequencies represented in the spectral envelope of whispered speech can be used by the listeners to match the voice pitch of the speaker. However, tone recognition of whispered speech in other tonal languages was typically found to be much lower than in Mandarin Chinese: ~40%

correct for Vietnamese (Miller 1961) and 20–45% correct for Thai (Abramson 1978). Since the duration and overall amplitude cues were poorly controlled in the whispered-speech studies, the degree to which the spectral-envelope cue contributed to tone recognition remains unclear.

2.2 Interaction Between the Temporal and Spectral Cues

The contributions of spectral and temporal information to tone recognition have been examined in several studies that used a vocoder technique simulating multi-channel cochlear implants. Cochlear implant simulation, using a noise-excited vocoder, typically involves dividing speech signals into spectral bands, extracting temporal envelopes from each of the bands as modulators, modulating wide-band noise spectrally limited by the same bandpass filters, and summing all the amplitude-modulated narrow-band noise to form a reconstructed signal (see Xu and Pfingst 2008). The temporal fine structure in the speech single is therefore replaced by noise in the bands. However, the band-specific temporal-envelope cue is well preserved. It is possible to control the spectral resolution of the output signal by varying the number of spectral bands and the amount of temporal details is typically controlled by varying the cut-off frequency of the lowpass filters that are used to extract the envelopes.

Fu et al. (1998) first reported that an increase from 50 to 500 Hz in the lowpass cut-off frequencies of the envelope extractors greatly improved tone recognition, which indicates that when spectral information is limited, tone recognition can be improved by an increase in temporal details. On the contrary, an increase of spectral bands from 1 to 4 did not seem to improve tone recognition. Xu et al. (2002), however, reported a significant effect of spectral channels when the number was varied from 1 to 12. In particular, Xu et al. (2002) found a trade-off between spectral and temporal cues: tone-recognition performance with higher spectral resolution but less detailed temporal envelope was equivalent to that with low spectral resolution but more detailed temporal envelope (Fig. 14.2).

Besides the trade-off relationship between temporal and spectral cues, Kong and Zeng (2006) observed complementary contribution between temporal periodicity cues and spectral cues for Mandarin tone recognition in quiet and noise conditions. They found that in quiet, tone-recognition performance of 8-band 50-Hz lowpass cutoff condition was worse than that of 1-band 500-Hz lowpass cutoff condition, but this pattern was reversed in noise conditions. This indicates that the coarse spectral information may not be very useful for tone recognition in quiet but is of great importance for perception in noise, because the temporal envelope cues might be more susceptible to noise compared to the spectral cues. A more recent study has suggested that temporal envelope and periodicity information (i.e., ≤500 Hz) within different frequency bands may have differential contribution to tone recognition. Yuen et al. (2007) showed that Cantonese tone recognition was significantly better when the listeners were provided with temporal-envelope information from the 2 higher frequency bands (1–2 kHz and 2–4 kHz) rather than with that of the two lower frequency bands (60–500 Hz and 500–1000 Hz).

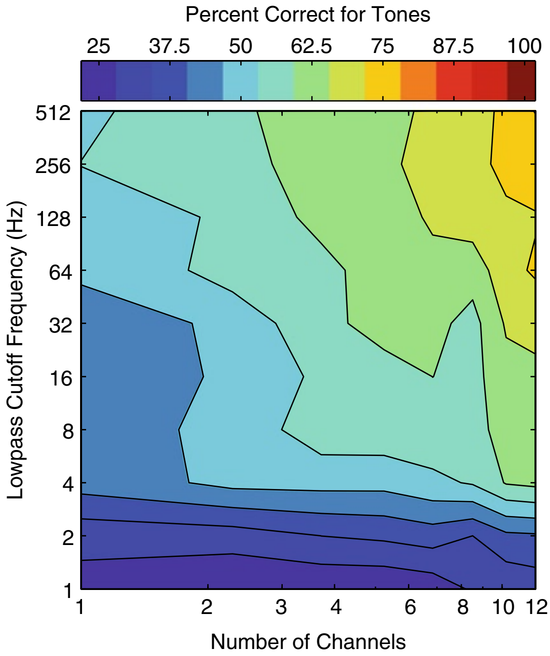

Fig. 14.2 Mean tone-recognition performance in the number-of-channels-versus-low-pass-cutoffs matrix. The data are plotted in a contour format in which the percent correct scores are represented by colors, as indicated by the color bar on the top. Adapted from Xu et al. (2002) with permission from the Acoustical Society of America

2.3 Relative Contributions of Temporal Envelope and Fine Structure

As early as 1988, Lin demonstrated that when the temporal fine structure of a broad-band signal exists, the temporal-envelope information has no influence on the recognition of Mandarin tone contrasts. The examination of the relative importance of temporal envelope and fine structure in multiple bands was made possible by using an “auditory chimera” signal-processing technique (Smith et al. 2002). This technique creates chimeric signals that have the temporal envelope of one tone and the fine structure of another tone (Xu and Pfingst 2003). Since the chimeric stimuli contain conflicting tonal information, the listeners’ response reveals whether it is the temporal fine structure or the envelopes that they depend on when making a tone judgment. Results of Xu and Pfingst (2003) showed that approximately 90% of the time the responses were consistent with the identity of fine structure of the chimeric stimuli.

The above findings have heralded new development of speech-processing strategies that have aimed at providing fine structure in the electrical stimulation in cochlear implants. Specifically, these are the HiResolution or HiResolution 120 strategies from the Advanced Bionics (Koch et al. 2004; Firszt et al. 2009) and the FSP strategy from Med El (Arnoldner et al. 2007). As reviewed below in Sect. 3, the emerging clinical data indicate little, if any, improvement in lexical-tone recognition using these new strategies. The reasons might be that the CI users are unable to use the fine structure information delivered in the electrical form or that perhaps the auditory system of a deaf individual has a reduced ability to process temporal

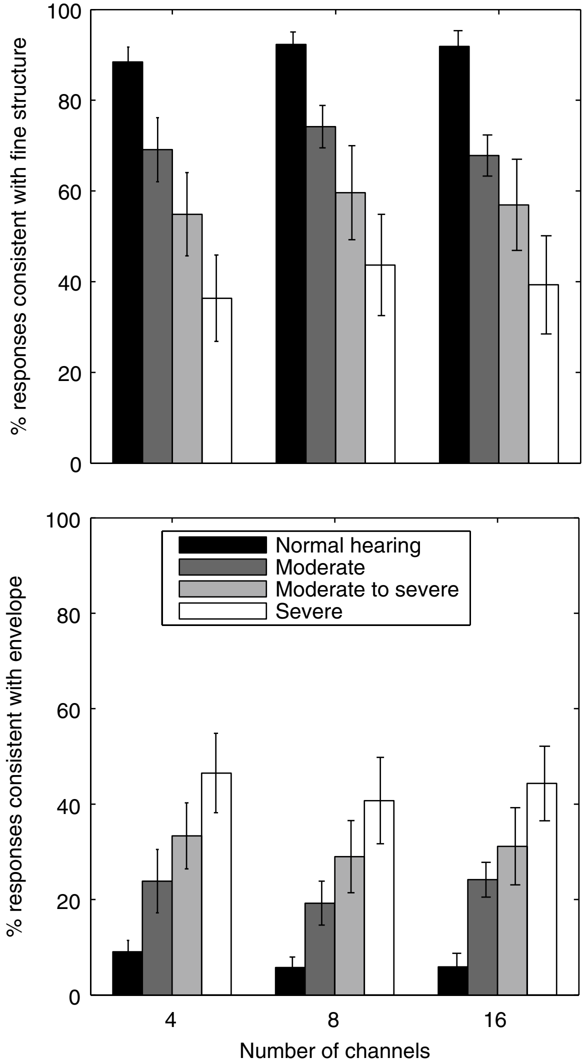

Fig. 14.3 Tone-recognition performance using chimeric stimuli in various number-of-channels conditions in normal-hearing listeners and patients with various degrees of sensorineural hearing loss. *Upper and lower panel* represent percentages of responses consistent with the fine structure and the envelope of the chimeric stimuli, respectively

fine-structure information (e.g., Lorenzi et al. 2006). Recently, Wang et al. (2010) used the auditory chimera technique and tested tone recognition in a group of patients with moderate to severe sensorineural hearing loss. Results clearly indicate that while normal-hearing listeners rely on fine structure for tone recognition, the hearing-impaired patients rely more on the temporal envelope for tone recognition as the hearing loss becomes more severe (Fig. 14.3).

Collectively, the literature suggests that F0 and its harmonic structures of the signal are the most dominant cues for tone recognition. In the absence of explicit F0, such as in CI stimulation or its vocoder simulation, temporal information, particularly the temporal envelopes presented in multiple channels that are equivalent to what are available in current CI technology, contributes to a moderate level (70 to 80% correct) of tone recognition.

3 Tone Recognition

3.1 Tone-Recognition Performance in CI Users

Tone-recognition performance in children with CIs who speak tonal languages is highly variable. Data from Mandarin-speaking as well as Cantonese-speaking children with CIs generally indicate that the prelingually deafened children with CIs have deficits in perceiving lexical tones. Wei et al. (2000) measured tone-recognition accuracy in 28 CI children (2–12 years of age) who speak Cantonese. Tone recognition improved significantly from pre-implantation, but performance plateaued at 65% correct after 2 years of training. Ciocca et al. (2002) tested tone recognition in 17 native Cantonese-speaking children aged between 4 to 9 years old. The subjects identified the tone in a monosyllabic target /ji/ in a two-alternative forced-choice test. The tone target was presented in a sentence median position. The subjects responded by pointing to the pictures that represented the auditory stimuli. Performance ranged from chance (i.e., 50% correct) to 61% correct for the 8 tone contrasts. Similar results have been reported by Lee et al. (2002) and Wong and Wong (2004) using either tone discrimination or tone identification paradigms.

Using a tone contrast test for Mandarin tone recognition (chance = 50% correct), Peng et al. (2004) reported a relatively higher average score of approximately 73% correct in a group of 30 pediatric Mandarin-speaking implantees aged from 6 to 12.6 years old. Among the 6 Mandarin Chinese tone contrasts, Peng et al. reported that the pediatric implantees could identify the tone contrasts that involve tone 4 better than other tone contrasts. Note that the researchers used a live voice presentation and did not control for the durations of the tone tokens.

A large-scale study was conducted recently on tone development in children with CIs who speak Mandarin Chinese (Xu et al. 2009b). The study tested tone recognition in 107 children aging from 2.4 to 16.2 years old. The study used a computerized two-alternative forced-choice paradigm in which the durations of the tone tokens were equalized. Performance of individuals varied considerably from chance to as high as a nearly perfect score (Fig. 14.4). Identification of the 6 tone contrasts did not show significant differences. The averaged tone-recognition accuracy of the group was 67% correct, significantly lower than the nearly perfect performance of the typically-developing, normal-hearing control group ($N = 112$). The children with CIs could perform relatively well whenever tone 1 was contrasted with other tones. Confusion matrix analysis also showed that tone 1 overall was the best recognized tone.

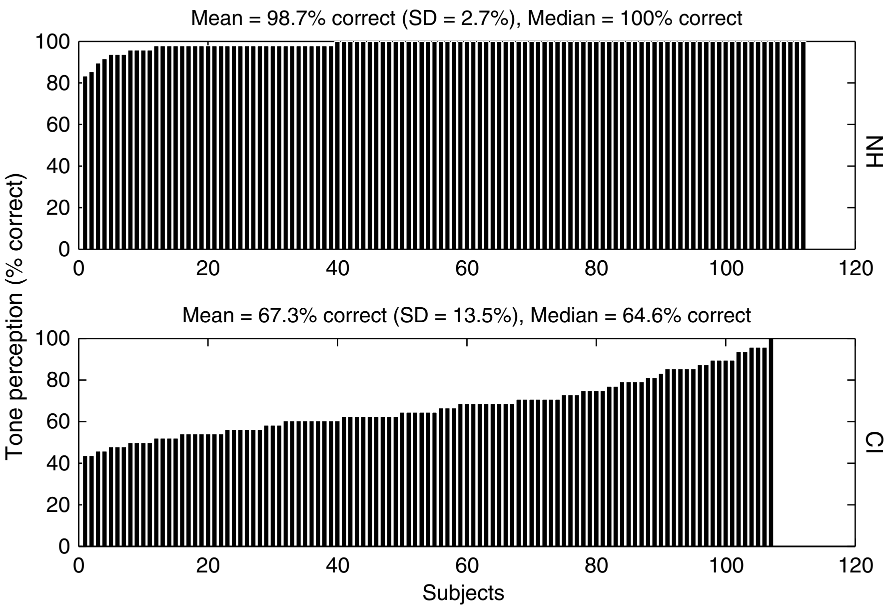

Fig. 14.4 Rank-ordered tone-recognition scores (% correct) by the 112 normal-hearing children (*upper panel*) and 107 children with CIs (*lower panel*)

3.2 Psychophysics and Tone Recognition

Lexical tone-recognition performance in CI users has been shown to correlate with a number of psychophysical measures including electrode discrimination, rate discrimination, gap detection, frequency discrimination, amplitude modulation detection, and amplitude modulation frequency discrimination thresholds. Wei et al. (2004) for example, measured pulse rate discrimination in individual electrodes in 5 CI subjects who speak Mandarin Chinese. Two standard rates (i.e., 100 and 200 pps) were chosen because they were in the range of the voice pitch. The average rate discrimination thresholds varied significantly in individuals ranging from 0.2 to 1.2 measured in Weber's fraction. The 5 CI users were also tested for tone recognition using various numbers of active electrodes. Despite the fact that the results varied largely between electrode conditions and implant subjects, Wei et al. (2004) found that tone-recognition scores using a full-electrode map (i.e., 20 electrodes) were highly correlated ($r = -0.97$) with the subjects' averaged rate discrimination thresholds across the electrode array. Wei et al. (2007) also compared tone recognition in 17 Mandarin-speaking implant users with their gap detection and frequency discrimination thresholds. They found that tone recognition in noise conditions showed stronger strength in correlation with the psychophysical measures. Luo et al. (2008) used a research interface to bypass the clinical speech processor to measure psychophysical performance. Amplitude modulation detection thresholds (AMDT) and amplitude modulation frequency discrimination thresholds (AMFDT) were measured in 10 Mandarin-speaking implant users at their middle electrodes with various stimulation levels. Results showed that the mean AMDTs (averaged for 20- or 100-Hz AM across different levels) and mean AMFDTs (averaged for the 50-Hz standard AM frequency across different levels) were significantly correlated with Mandarin Chinese tone, consonant, and sentence recognition scores, but not with vowel recognition scores. Their results further confirmed the importance of temporal-envelope cues for Chinese speech recognition in CI users.

3.3 The Effects of Frequency-Place Mismatch on Tone Recognition

In the normal cochlea, acoustic signals of different frequencies stimulate corresponding places on the basilar membrane in a tonotopic fashion. In CI systems, a number of forms of frequency-place mismatch may occur as a result of the pathology of hearing loss, shallow insertion of the electrode, or frequency mapping of the device. Localized losses of auditory neurons can result in "holes" in hearing and elevate electrical thresholds of the corresponding electrodes. The increased signal level will likely result in spread of electric current to neural fibers that are not intended to be activated, producing frequency warping around the "holes" in the cochlea (Shannon et al. 2002). Frequency-place mismatch can also take place in shallow insertion conditions that result in an overall shift of the spectrum. Consider the case where the implant electrode array is not fully inserted into the cochlea so that the location of the electrode array does not match the analysis bands. Typically, the output of a low frequency analysis band is delivered to the electrode at a higher frequency place, resulting in a basal shift of the spectrum. Matching the analysis bands to the location of the electrode array nonetheless eliminates frequency coverage especially in the low frequency region. Additionally, because of the limited length of the cochlear implant electrode array, the frequency range stimulated by a cochlear implant does not necessarily cover the entire speech spectrum. As a consequence, frequency compression is another commonly encountered case of frequency-place mismatch. Clinically used maps usually compressively allocate a wider frequency range, sufficient for speech understanding to electrodes that cover a narrower tonotopic length, regardless of the position of the electrode array.

There is a consensus in the literature that suggests a detrimental effect of frequency-place mismatch on consonant, vowel, and sentence recognition in English (see Dorman et al. 1997; Pfingst et al. 2001; Baskent and Shannon 2006). Basal shift also shows a systematic effect on English consonant confusion (Zhou et al. 2010). The effects of basal spectral shift and frequency compression on lexical-tone recognition were examined by Zhou and Xu (2008b). In the study, a noise-excited vocoder was used to simulate a cochlear implant with varying insertion depths. Speech envelopes were delivered to carriers of higher frequencies to simulate basal spectral shift of 1–7 mm. Zhou and Xu (2008b) found that tone recognition was much more resistant to the basal spectral shift compared to English phoneme and sentence recognition. The detrimental effects of basal shift did not show until the carriers were shifted to

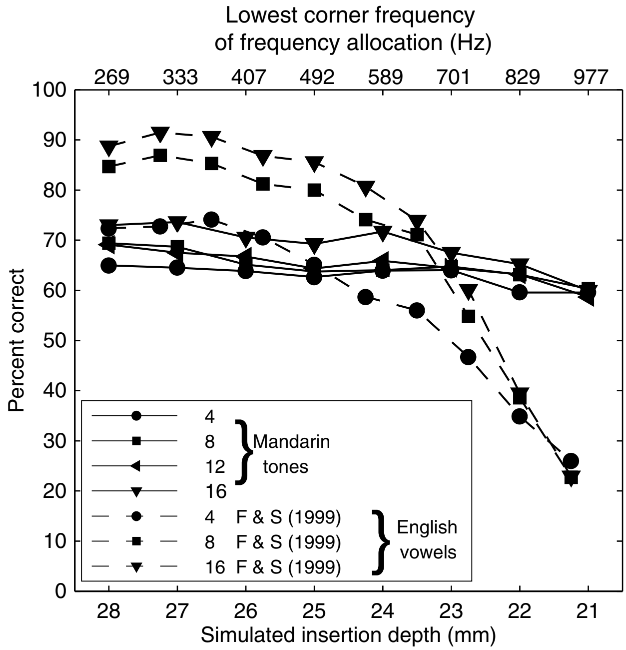

Fig. 14.5 Tone-recognition performance as a function of simulated insertion depth of cochlear implants (*CIs*). Tone-recognition performance is plotted for 4, 8, 12, and 16 channel conditions in *solid lines* with different symbols. The lowest corner frequency of frequency allocations for the carriers is noted for each simulated insertion depth. Data of vowel recognition from Fu and Shannon (abbreviated as F & S in the legend) (1999) are replotted. Simulated insertion depth of 28 mm corresponds to a full insertion or tonotopically matched condition. Adapted from Zhou and Xu (2008b) with permission from Elsevier

almost 2 octaves higher. A 7-mm basal shift of the spectrum only caused tone-recognition performance to decrease from the unshifted condition by approximately 10 percentage points. Compared to the vowel recognition scores measured from similar experimental conditions (Fu and Shannon 1999), the effects of spectral shift on tone recognition appeared to be much smaller (Fig. 14.5).

Zhou and Xu (2008b) also reported the effect of frequency compression on Mandarin tone recognition. Compression of 3 or 5 mm at both frequency ends on the basilar membrane of the cochlea (Greenwood 1990) produced better tone-recognition performance than that without compression. This is consistent with the findings by Baskent and Shannon (2003, 2005) that for English phoneme recognition, in shallow insertion conditions, a moderate amount of compression is better than tonotopic matching with low frequency truncation. The Zhou and Xu study (2008b) suggests that wider frequency allocation that includes low frequency information critical for pitch perception may benefit tone recognition. However, the degree of compression should be controlled so that the effects of frequency mismatch will not cancel out the benefit of frequency coverage.

Research data from a limited number of CI users are consistent with the vocoder study discussed above in that frequency-place mismatch affects Mandarin tone recognition much less than word recognition. In such a study, Liu et al. (2004) tested the effects of frequency shift and compression on Mandarin tone and word recognition in 6 prelingually deafened children fit with Nucleus CI24. The frequency range was kept constant, while electrodes were selectively turned off, to create frequency shift and compression combined conditions. Their results suggested that as long as a sufficient number of electrodes were activated, selection of the stimulation sites did not seem to affect tone recognition. It was not possible however to separate the effects of frequency shift and frequency compression in Liu et al. (2004), because the same frequency range was used in the experiments.

3.4 Demographic Factors Contributing to Tone Recognition

Many studies have tried to explain the variable performance in the CI children in relation to their demographic variables such as age at implantation and the experience with the devices. For example, Lee et al. (2002) reported that tone-recognition performance was related to the duration of CI use and age at implantation. Han et al. (2009) found a consistent relationship between tone recognition performance and age at implantation. Nonetheless, Peng et al. (2004) and Wong and Wong (2004) did not find significant correlation between tone-recognition performance and any of the potential predictive variables. The limited sample sizes in these studies may explain the discrepancies.

Xu et al. (2009b) collected a much larger sample size and were able to study a number of other predictor variables as potential contributors to tone recognition in addition to the demographic variables. These predictors included family variables (such as family size and household income), cochlear implant variables (such as implant type and speech processing strategy), and educational variables (such as communication mode and duration of speech therapy). All predictor variables were entered step-wise into a linear regression model for analysis. The regression analysis, however, showed that only the demographic variables were the significant predictors for tone recognition. The study showed that jointly age at implantation and duration of CI use were significant predictors to tone recognition. Xu et al. (2009b) reported that duration of CI use was a stronger predictor, because it had a significant marginal relationship with tone recognition. Age at implantation alone, however, could not explain a significant amount of the total variance in tone recognition. Therefore, the marginal relationship was not significant. When both variables were entered into the regression model, in the presence of duration of CI use, age at implantation explained a significant amount of unique variance in tone recognition. Jointly, they explained approximately 50% of the total variance for tone-recognition outcome. The results indicate that although the performance of perception is bound to improve as the experience of using the device increases, early implantation may facilitate this improvement.

3.5 CI Stimulation Features and Tone Recognition

Although Xu et al. (2009b) did not find predictor variables associated with tone-recognition performance other than age at implantation and duration of CI use, several studies have indicated that CI tone recognition is related to speech-processing strategies. Fu et al. (2004) showed that ACE and CIS were better than SPEAK for Mandarin tone recognition. They speculated that the stimulation rate in the ACE and CIS strategies are higher than the SPEAK strategy (typically at 250 pps), providing higher temporal resolution to better encode the voice pitch. On the other hand, Barry et al. (2002) found similar Cantonese tone recognition between children who used ACE and SPEAK strategies. These contradictory findings might be the result of differences in stimulus, subject, and experience. For example, in Fu et al. study, all subjects had over 3 years of experience with the high-rate ACE strategy as opposed to limited experience with the SPEAK strategy.

Vandali et al. (2005) tested a novel speech-processing strategy, namely MEM (i.e., multi-channel envelope modulation) that was designed to enhance coding of F0 periodicity cues in the speech signal. In this strategy, the low frequency (80–400 Hz) envelope of the broadband signal, which contains F0 periodicity information for voiced/periodic signals, was used to modulate the envelope of each analysis band of the ACE strategy. CI subjects performed significantly better in pitch-ranking tests using the novel strategy than with ACE and CIS. Based on the encouraging results, Wong et al. (2008) evaluated the MEM strategy in a group of Cantonese-speaking adult CI users. Although tone recognition was not measured in the study, no difference was found in sentence recognition using the Cantonese version of HINT (Wong and Soli 2005) between ACE and MEM strategies.

Based on the psychophysical evidence that simultaneous stimulation of 2 adjacent electrodes results in a pitch percept that is between those elicited by stimulation of the 2 electrodes individually (see Bonham and Litvak 2008 for a review), Advanced Bionics introduced HiResolution with Fidelity 120 (HiRes 120) in 2006. Han et al. (2009) studied whether HiRes 120, which presumably provides much finer spectral resolution than the traditional strategies, would benefit lexical tone recognition. Twenty Mandarin-speaking children who were originally fitted with HiRes were experimentally switched to the new HiRes120 for a period of 6 months. Tone-recognition performance with the HiRes120 was comparable to the baseline performance with the HiRes strategy as a group. Some benefits were observed for approximately one third of the individuals after a period of either 3 or 6 months after the strategy conversion. Similar results were obtained in a separate study on HiRes 120 by Chang et al. (2009).

In an effort to enhance pitch coding, Med El recently launched a Fine Structure Coding Strategy (FSP) in which "packets" of pulses that represent zero-crossings of the acoustical signals in the low frequency bands are delivered to the apical electrodes (Arnoldner et al. 2007; Riss et al. 2009). Schatzer et al. (2010) compared tone-identification performance with the FSP strategy and the traditional CIS strategy in 12 Cantonese-speaking adult CI users. Their preliminary results showed no significant differences between the two strategies in an acute experiment.

There are a number of other speech-processing strategies that were proposed to enhance tonal-language recognition with CIs and were tested in acoustic simulations. Lan et al. (2004) proposed to use F0 as a carrier to replace the fixed-rate carrier in the standard CIS strategy. Nie et al. (2005) extracted slowly varying frequency modulation to encode the temporal fine structure. Luo and Fu (2004) modified the temporal envelope as well as the modulation depth of the periodicity fluctuation in local channels to better resemble the F0 contour. Alternatively, the overall amplitude contour was adjusted based on the F0 contour before the vocoder processing. Yuan et al. (2009) attempted to replace the temporal envelopes in the high frequency bands by a sinusoid with frequency equal to the F0 and found that Cantonese tone-recognition performance significantly improved. All these manipulations in signal processing have not been implemented in CI processors.

3.6 The Relation Between Musical and Lexical Pitch Perception

Because of inadequate representation of pitch information in the cochlear implant systems, implant users also have difficulties in perceiving musical pitch (see McDermott, Chap. 13). Studies indicate that postlingually deafened implant users consistently show impairment in identifying familiar songs without rhythmic cues (e.g., Gfeller et al. 2002, 2007; Leal et al. 2003). Many of them reported that the enjoyment of listening to music declines substantially after implantation (Lassaletta et al. 2007). A handful of studies specifically examined pitch-discrimination ability in CI users that is directly linked to their ability to perceive music. Fujita and Ito (1999) reported that the pitch-ranking thresholds measured from 8 CI users fell in a wide range of 4 semitones to 2 octaves. Looi et al. (2008) reported that CI subjects were unable to rank pitches that were a quarter-octave (i.e., 3 semitones) apart. They were only able to rank pitches that were a half-octave and one-octave apart 64% and 68% of the time correctly, respectively. Similar results of CI users' performance in pitch ranking were reported by Sucher and McDermott (2007). Other studies used adaptive procedures and reported pitch-discrimination thresholds that were in the range of 1–12 semitones (Gfeller et al. 2002; Nimmons et al. 2008; Kang et al. 2009).

Based on the notion that music appreciation and tone recognition both involve pitch perception, intuitively these two aspects of pitch perception should correlate with each other. Such a relationship was reported by Wang et al. (2010) who examined the mechanisms of musical and lexical tone recognition using CIs. Using a novel method to measure music perception that had several advantages over the traditional pitch ranking or familiar melody tests, they found that the discrimination thresholds of the CI subjects were highly variable, ranging from 0.65 to 19 semitones and were significantly worse than those of the normal-hearing controls (Fig. 14.6, left panel). Tone-recognition scores from the CI subjects ranged from 12.5% to 86.8% correct (Fig. 14.6, middle panel). More importantly, a highly significant negative

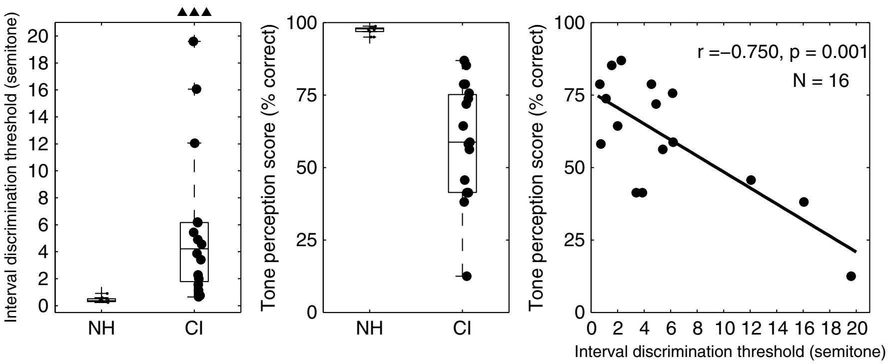

Fig. 14.6 Musical and lexical tone recognition performance. *Left panel*: Box plot of the averaged pitch interval discrimination thresholds in 10 normal-hearing (NH) and 19 CI subjects. The 3 triangles plotted at the top represent the 3 CI subjects who could not perform the interval discrimination test even at a ΔF0 of 2 octaves. *Middle panel*: Box plot of the Mandarin-Chinese tone-recognition scores for normal-hearing ($N=10$) and CI ($N=19$) subjects. *Right panel*: Correlation between tone-recognition scores and averaged pitch interval discrimination thresholds in CI subjects. Each symbol represents one subject with a CI. The *solid line* represents the linear fit of the data

correlation was shown between the pitch interval discrimination thresholds and the tone-recognition performance in the CI subjects ($r=-0.75$, $N=16$, $p<0.01$) (Fig. 14.6, right panel).

Wang et al. (2010) suggested that the strong correlation between the musical interval discrimination threshold and the tone-recognition performance indicated a shared mechanism in electric stimulation for musical and voice pitch perception. Musical pitch can be perceived via temporal patterns in amplitude modulation over restricted and relatively low modulation frequencies (e.g., McKay et al. 1994). Likewise, lexical tone information is supported by the periodicity in the temporal envelopes. Although they speculated that the temporal pitch coding underlies the common mechanisms for musical pitch and lexical tone recognition, it should not be ruled out that pitch perception can be realized through different excitation patterns of the electrical stimulation via a more place coding mechanism (e.g., Pretorius and Hanekom 2008).

4 Tone Production

4.1 Tone Development in Normal-Hearing Children

There is a substantial body of literature related to normal-hearing children's phonological development, but only a few on the tone acquisition of children who speak tonal languages. The earliest account of tone acquisition in normal-hearing children was from Chao (1951), who reported that his granddaughter acquired tones very early, and her isolated tones of stressed syllables were practically the same as in standard Mandarin (cited by Li and Thompson 1977 and Tse 1978). Li and Thompson (1977) studied tone acquisition in 17 Mandarin-speaking children ranging in age from 1.5 to 3.0 years old. Although no exact time frame for tone development was provided and the finding was descriptive in nature, Li and Thompson (1977) concluded that tone acquisition is accomplished very early in age. Tone acquisition takes place within a relatively short period of time and is well in advance of the mastery of segmentals. This notion was first suggested in a longitudinal case study (Tse 1978). Tse reported that his son's tone acquisition peaked between 14 and 22 months of age. More recently, Zhu and Dodd (2000) studied phonological acquisition in a large group of Mandarin-speaking children. Tone production in 21 children of their youngest subgroup (1.5–2.0) was reported fairly accurate. Both Li and Thompson (1977) and Zhu and Dodd (2000) used a picture-naming procedure to elicit production. One or two experienced Mandarin-speaking phoneticians transcribed the tone production as a measure for production performance. Wong et al. (2005) used 10 Mandarin-speaking adults to judge the tone production of 13 3-year-old Mandarin-speaking children residing in the United States. They found that the children had not fully mastered the production of the 4 Mandarin Chinese tones in monosyllabic words. From the limited data on tone acquisition, there is no clear agreement on the exact time frame that tone acquisition is complete.

4.2 Tone-Production Performance in CI Children

Accompanying the difficulties in perceiving tones, prelingually deafened implant users have also shown challenges in tone production, probably as a result of distorted auditory input of the tone targets (Xu et al. 2004, 2010). A converging finding of tone-production performance in those children is that, like tone recognition, there is a great individual difference among the users (Peng et al. 2004; Xu et al. 2004, 2009b; Han et al. 2007; Lee et al. 2010). Peng et al. (2004) reported tone-production performance in a group of 30 prelingually deafened, Mandarin-speaking children with CIs. The average percent-correct score for the group was 62%. The lowest score was about 20% correct, while 2 of the children scored nearly perfect. Peng et al. found that the children's performance was correlated with their age at implantation but was not related to the experience of the device use. More recently, Lee et al. (2010) evaluated Cantonese tone production in a longitudinal study. Their results indicated that children with earlier implantation (<4 years old) achieved more effective acquisition of lexical tones than those who received CIs after 4 years of age.

Tone production of Mandarin-speaking prelingually deafened children with CIs was further explored by Han et al. (2007). Native adult listeners were not asked to give any subjective judgments, but rather, they were asked to identify the tones they heard out of 4 possible choices (chance = 25% correct). The averaged production accuracy of the CI group was 48% correct, significantly lower than that of the

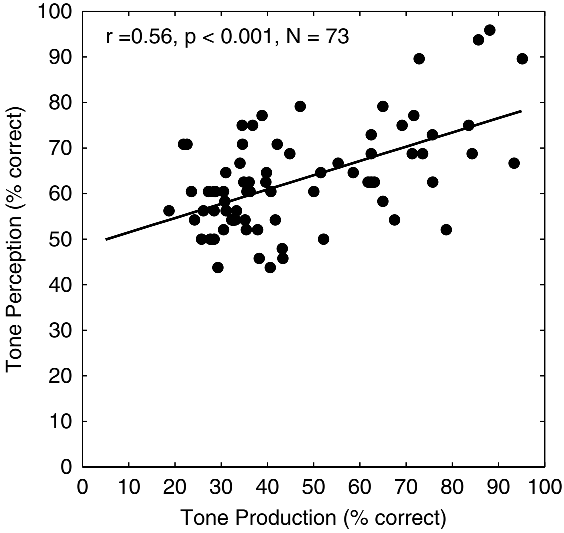

Fig. 14.7 Correlation between tone recognition and production performance in CI children. Tone-production scores were those evaluated by the normal-hearing adult listeners. Each symbol represents 1 subject with a CI. The solid line represents the linear fit of the data

age-matching normal-hearing control group (78% correct). The CI subjects in the study had particular difficulties in producing tone 2 followed by tones 3 and 4. The normal-hearing, native Mandarin-speaking adult listeners often perceived the CI children's intended contour tones as a flat tone (i.e., tone 1). In a follow-up study with more subjects, Xu et al. (2009b) also observed that the implanted children seemed to have particular difficulties with producing tone 2. This is consistent with the report by Peng et al. (2007), which revealed deficits in the production of rising intonation in native English-speaking pediatric CI users. Nonetheless, the difficulty of producing the rising tone 2 cannot seem to be attributed to the ability of perceiving this tone. It is more likely caused by the efforts demanded in producing the rising pitch. The ability of perception, similarly, cannot fully explain the difficulties of the children with CIs had in producing tone 3, which had the most sophisticated pitch contour. Although perception did not seem to predict the children's production performance for each tone, the perception accuracy predicted their overall production performance, and vice versa (Fig. 14.7). This correlation suggested that the development in tone recognition and production in children with CIs in general are two related aspects of language that facilitate each other (Xu et al. 2010).

Studies that examined the acoustic properties of the tone production from children with CIs revealed interesting findings. One of the approaches for the acoustic analysis is based on the distributions of the onsets and offsets of the F0 contours (Barry and Blamey 2004). Such distributions form the tonal ellipses, the overlap of which can be quantified based on Signal Detection Theory (Green and Swets 1966). Figure 14.8 shows an example of the tonal ellipses from 4 normal-hearing, native Mandarin-speaking children and 4 prelingually deafened children with CIs. The tonal ellipses reflect two components of the acoustic properties, which are the variability in the F0 use for one tone (i.e., ellipse size) and the overall F0 span that measures the tonal area (i.e., ellipse overlap).

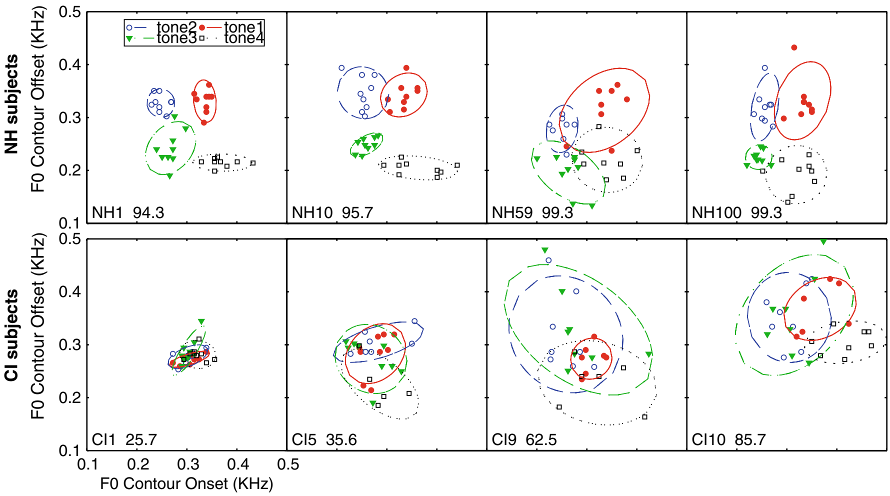

Fig. 14.8 Tonal ellipses based on distributions of F0 onsets and offsets of F0 contours produced by children. The 4 subjects from normal-hearing group (*upper panels*) and 4 from prelingually-deafened CI group (*lower panels*) were randomly selected from those with tone-production scores (judged by the native adult listeners) in the ranges of 0–25, 25–50, 50–75, and 75–100 percentiles of each respective group. Each data point represents a pair of F0 onset-offset value of a monosyllabic word. Symbols in different colors and styles represent different tones, as indicated by the legend in the *top left-hand panel*. Each *ellipse* encompasses 95% of the data points for each tone

The F0 variability, as assessed by the tonal ellipse analysis mentioned above, shows an age-related function in the typically-developing normal-hearing children. Zhou and Xu (2008a) showed that the F0 use of individual tones by the normal-hearing group presents a more confined pattern with age (i.e., ≥6 years of age). Such a development, however, was not found in the CI group with their duration of device use. That means that the F0 variability does not improve even with the accumulated experience with the device (Zhou and Xu 2008a). Xu et al. (2009b) confirmed that a relatively older normal-hearing group (≥6 years of age) have a significantly smaller variability of F0 use for a particular tone than that of the CI group. The F0 span or tonal area that measures the overall F0 range has always been reported to be much larger in the normal-hearing group than the CI children group. Zhou and Xu (2008a) indicated that even though the younger normal-hearing children have not learned to use a certain F0 range consistently for producing individual tones, their large F0 span or tonal area compensates for such variability; thus their tone production remains more differentiable than the CI children.

In addition to the acoustic analysis, artificial neural networks have been used to evaluate tone production in tonal-language speakers (e.g., Xu et al. 2006, 2007; Zhou et al. 2008). A feed-forward multilayer perceptron has been used to recognize the tones produced by a group of Mandarin-speaking children with normal hearing (Xu et al. 2007). The neural network provides direct classification results, from which the recognition percent-correct scores as well as tone confusion matrices can

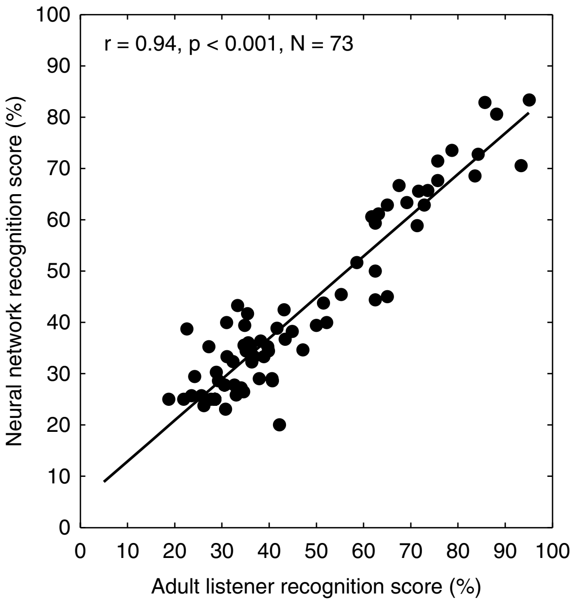

Fig. 14.9 Correlation of tone-production performance by CI subjects evaluated by a neural network and by normal-hearing adult listeners. Each symbol represents 1 subject with a CI. The *solid line* represents the linear fit of the data

be generated. The error patterns of the neural network were remarkably similar to that of the human listeners (Xu et al. 2007; Zhou et al. 2008). Xu et al. (2009b) tested tone production in a group of 73 prelingually deaf children with CIs and found that the neural-network data correlated strongly with the perceptual judgments by adult native listeners ($r=0.94$, $N=73$, $p<0.001$) (Fig. 14.9).

5 Vocal Singing with CIs

As discussed in 3.6, musical pitch perception and lexical tone recognition may share a similar mechanism in electric hearing. As a consequence of poor pitch perception, prelingually deafened children with CIs have demonstrated poor development in vocal singing similar to that in tone production. Nakata et al. (2005) studied vocal singing in 12 congenitally deafened children (4.9–10.3 years of age) who had received CIs. Children with CIs could sing familiar songs from their memory, although the pitch patterns were largely unrelated to the direction of pitch patterns in the target songs (Nakata et al. 2005). Xu et al. (2009a) further explored vocal singing in 7 prelingually deafened children with CIs (age: 5.4–12.3 years old). The control group consisted of 14 normal-hearing children (age: 4.1–8.0 years old). The production of music pitch was evaluated acoustically. In the study, five metrics were developed based on the acoustic analysis of the F0 contours of the sung notes.

The five metrics included (1) F0 contour direction of the adjacent notes, (2) F0 compression ratio of the entire song, (3) mean deviation of the normalized F0 across the notes, (4) mean deviation of the pitch intervals, and (5) standard deviation of the note duration differences. Compared to the normal-hearing children, the CI group performed significantly poorer in the first four metrics that assessed the pitch-based performance of vocal singing. Similar to tone production, a large individual difference was observed. Singing of the CI children tended to be monotonic, compressed in the F0 range, and largely unrelated to the pitch contour of the target songs. No significant differences were seen between the two groups in the rhythm-based measure (i.e., the fifth metric). Current CI systems can faithfully deliver rhythmic information. Patients with CIs have been reported to perform at a level similar to normal-hearing subjects in rhythmic perception tasks (e.g., Gfeller et al. 1997; Kong et al. 2004; see also McDermott, Chap. 13). Thus, as a result of the preserved rhythmic information, the rhythmic aspect of singing in implanted children might not be significantly affected.

6 Summary

The primary acoustic cue for tone recognition is temporal or spectral fine structure of the signal. When the primary cue is not available, as in electric hearing, the temporal cues such as the amplitude contour, periodicity in the amplitude modulation patterns, or duration serve as secondary cues for tone recognition. Tone recognition is less susceptible to the spectral distortions because of the contribution of the secondary temporal cues.

Tone development in children with CIs demonstrates different patterns than the normal-hearing children, as revealed by the acoustic properties of their tone production and the error patterns of tone recognition. Although the overall performance of tone recognition is correlated with that of tone production in children with CIs, their error patterns of perception and production are not associated. Earlier implantation and more experience with the implant device predict better tone recognition, but only younger age at implantation predicts better tone production in children with CIs. Because of poor pitch encoding in the contemporary CI systems, prelingually deafened CI users who speak tonal languages generally have poor tone production and vocal singing. Some tonal-language users can probably maintain a fairly high level of speech perception using contextual cues when tone information is sparse. The long-term impact of poor tone recognition on tonal-language development remains to be explored in pediatric CI users.

Acknowledgements We thank Heather Schultz and Marisol Gliatas for the technical support during the preparation of the manuscript. The work was supported in part by NIH NIDCD Grants R03-DC006161, R15-DC009504, and F31-DC009919.

References

Abramson, A. S. (1978). Static and dynamic acoustic cues in distinctive tone. Language Speech, 21, 319–325.

Arnoldner, C., Riss, D., Brunner, M., Durisin, M., Baumgartner, W.-D., & Hamzavi, J.-S. (2007). Speech and music perception with the new fine structure speech coding strategy: preliminary results. *Acta Oto-Laryngologica*, 127, 1298–1303.

Barry, J. G., & Blamey, P. J. (2004). The acoustic analysis of tone differentiation as a means for assessing tone production in speakers of Cantonese. *Journal of the Acoustical Society of America*, 116, 1739–1748.

Barry, J. G., Blamey, P. J., Martin, L. F. A., Lee, K. Y. S., Tang, T., Ming, Y. Y., & Van Hasselt, C. A. (2002). Tone discrimination in Cantonese-speaking children using a cochlear implant. *Clinical Linguistics & Phonetics*, 16, 79–99.

Baskent, D., & Shannon, R. V. (2003). Speech recognition under conditions of frequency-place compression and expansion. *Journal of the Acoustical Society of America*, 113, 2064–2076.

Baskent, D., & Shannon, R. V. (2005). Interactions between cochlear implant electrode insertion depth and frequency-place mapping. *Journal of the Acoustical Society of America*, 117, 1405–1416.

Baskent, D., & Shannon, R. V. (2006). Frequency transposition around dead regions simulated with a noiseband vocoder. *Journal of the Acoustical Society of America*, 119, 1156–1163.

Bonham, B. H., & Litvak, L. M. (2008). Current focusing and steering: modeling, physiology, and psychophysics. *Hearing Research*, 242, 141–153.

Chang, Y. T., Yang, H. M., Lin, Y. H., Liu S. H., & Wu, J. L. (2009). Tone discrimination and speech perception benefit in Mandarin speaking children fit with HiRes fidelity 120 sound processing. *Otology & Neurotology*, 30, 750–757.

Chao, Y. R. (1951). The Cantian idiolect: an analysis of the Chinese spoken by a twenty-eight-month-old child. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development*. New York: Holt, Rinehart and Winston, Inc.

Ciocca, V., Francis, A. L., Aisha, R., & Wong, L. (2002). The perception of Cantonese lexical tones by early-deafened cochlear implantees. *Journal of the Acoustical Society of America*, 111, 2250–2256.

Dorman, M. F., Loizou, P. C., & Rainey D. (1997). Simulating the effect of cochlear implant electrode insertion depth on speech understating. *Journal of the Acoustical Society of America*, 102, 2993–2996.

Duanmu, S. (2000). *The phonology of standard Chinese.* Oxford: Oxford University Press.

Firszt, J. B., Holden, L. K., Reeder, R. M., & Skinner, M. W. (2009). Speech recognition in cochlear implant recipients: comparisons of standard HiRes and HiRes120 sound processing. *Otology & Neurotology*, 30, 146–152

Fu, Q.-J., & Shannon, R. V. (1999). Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. *Journal of the Acoustical Society of America*, 105, 1889–1990.

Fu, Q.-J., & Zeng, F.-G. (2000). Identification of temporal envelope cues in Chinese tone recognition. *Asia Pacific Journal of Speech, Language and Hearing*, 5, 45–57.

Fu, Q.-J., Zeng, F.-G., Shannon, R. V., & Soli, S. D. (1998). Importance of tonal envelope cues in Chinese speech recognition. *Journal of the Acoustical Society of America*, 104, 505–510.

Fu, Q.-J., Hsu, C. J., & Horng, M. J. (2004). Effects of speech processing strategy on Chinese tone recognition by nucleus-24 cochlear implant users. *Ear and Hearing*, 25, 501–508.

Fujita, S., & Ito, J. (1999). Ability of nucleus cochlear implantees to recognize music. *Annals of Otology, Rhinology & Laryngology*, 108, 634–640.

Gfeller, K., Woodworth, G., Robin, D. A., Witt, S., & Knutson, J. F. (1997). Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users. *Ear and Hearing*, 18, 252–260.

Gfeller, K., Turner, C., Mehr, M., Woodworth, G., Fearn, R., Knutson, J. F., Witt, S., & Stordahl, J. (2002). Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implant International*, 3, 29–53.

Gfeller, K., Turner, C., Oleson, J., Zhang, X. Y., Gantz, B., Froman, R., & Olszewski, C. (2007). Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. *Ear and Hearing*, 28, 412–23.

Gottfried, T. L., & Suiter, T. L. (1997). Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones. *Journal of Phonetics*, 25, 207–231.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Greenwood, D. D. (1990). A cochlear frequency-position function for several species-29 years later. *Journal of the Acoustical Society of America*, 87, 2952–2605.

Han, D., Zhou, N., Li, Y., Chen, X., Zhao, X., & Xu, L. (2007). Tone production of Mandarin Chinese speaking children with cochlear implants. *International Journal of Pediatric Otorhinolaryngology*, 71, 875–880.

Han, D., Liu, B., Zhou, N., Chen, X., Kong, Y., Liu, H., Zheng, Y., & Xu, L. (2009). Lexical Tone recognition with HiResolution® 120 Speech-Processing Strategy in Mandarin-Speaking Children. *Ear and Hearing*, 30, 169–177.

Howie, J. (1976). *An acoustic study of Mandarin tones and vowels*. London: Cambridge University Press.

Kang, R., Nimmons, G. L., Drennan, W., Longnion, J., Ruffin, C., Nie, K., Won, J. H., Worman, T., Yueh, B., & Rubinstein, J. (2009). Development and validation of the University of Washington Clinical Assessment of Music Perception test. *Ear and Hearing*, 30, 411–418.

Koch, D. B., Osberger, M. J., Segel, P., & Kessler, D. (2004). HiResolution™ and conventional sound processing in the HiResolution™ bionic ear: using appropriate outcome measures to assess speech recognition ability. *Audiology & Neurotology*, 9, 214–223.

Kong, Y. Y., & Zeng, F. G. (2006). Temporal and spectral cues in Mandarin tone recognition. *Journal of the Acoustical Society of America*, 120(5), 2830–2840.

Kong, Y. Y. Cruz, R., Jones, J. A., & Zeng, F. G. (2004). Music perception with temporal cues in acoustic and electric hearing. *Ear and Hearing*, 25, 173–185.

Kuo, Y.-C., Rosen, S., & Faulkner, A. (2008). Acoustic cues to tonal contrasts in Mandarin: Implications for cochlear implants. *Journal of the Acoustical Society of America*, 123(5), 2815–2824.

Lan, N., Nie, K., Gao, S., & Zeng, F. G. (2004). A novel speech-processing strategy incorporating tonal information for cochlear implants. *IEEE Transactions on Biomedical Engineering*, 51, 752–60.

Lassaletta, L., Castro, A., Bastarrica, M., Perez-Mora, R., Madero, R., de Sarria, J., & Gavilan, J. (2007). Does music perception have an impact on quality of life following cochlear implantation? *Acta Oto-Laryngologica*, 127, 682–686.

Leal, M. C., Shin, Y. J., Laborde, M. L., Calmels, M. N., Verges, S., Lugardon, S., Andrieu, S., Deguine, O., & Fraysse, B. (2003). Music perception in adult cochlear implant recipients. *Acta Oto-Laryngologica*, 123, 826–835.

Lee, C.-Y. (2009). Identifying isolated, multispeaker Mandarin tones from brief acoustic input: A perceptual and acoustic study. *Journal of the Acoustical Society of America*, 125, 1125–1137.

Lee, C.-Y., & Hung, T.-H. (2008). Identification of Mandarin tones by English-speaking musicians and nonmusicians. *Journal of the Acoustical Society of America*, 5, 3235–3248.

Lee, K. Y. S., van Hasselt, C. A., Chiu, S. N., & Cheung, D. M. C. (2002). Cantonese tone recognition ability of cochlear implant children in comparison with normal-hearing children. *International Journal of Pediatric Otorhinolaryngology*, 63, 137–147.

Lee, K. Y. S., van Hasselt, C. A., & Tong, M. C. F. (2010). Age sensitivity in the acquisition of lexical tone production: evidence from children with profound congenital hearing impairment after cochlear implantation. *Annals of Otology, Rhinology & Laryngology*, 119, 258–265.

Li, C. N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, 4, 185–199.

Liang, Z. A. (1963). Tonal discrimination of Mandarin Chinese. *Acta Physiologica Sinica* 26, 85–91.

Lin, M.-C. (1988). The acoustic characteristics and perceptual cues of tones in standard Chinese. *Zhongguo Yuwen*, 204, 182–193.

Liu, S., & Samuel, A. G. (2004). Perception of Mandarin lexical tones when f0 information is neutralized. *Language and Speech*, 47, 109–138.

Liu, T.-C., Chen, H.-P., & Lin, H.-C. (2004) Effects of limiting the number of active electrodes on mandarin tone perception in young children using cochlear implants. *Acta Oto-Laryngologica*, 124, 1149–1154.

Loizou P.C. (2006). Speech processing in vocoder-centric cochlear implants. *Advances in Otorhinolaryngology*, 64, 109–143.

Looi, V., McDermott, H., McKay, C., & Hickson, L. (2008). Music perception of cochlear implant users compared with that of hearing aid users. *Ear and Hearing*, 29, 421–434.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006) Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Science*, 103, 18866–18869.

Luo, C., & Wang, J. (1981). *Putong yuyinxue gangyao* [*Outline of general phonetics*], new ed. Beijing: Shangwu Yinshuguan.

Luo, X., & Fu, Q. (2004). Enhancing Chinese tone recognition by manipulating amplitude envelope: Implications for cochlear implant. *Journal of the Acoustical Society of America*, 116, 3659–3667.

Luo, X., Fu, Q.-J., Wei, C.-G., & Cao, K.-L. (2008). Speech recognition and temporal amplitude modulation processing by Mandarin-speaking cochlear implant users. *Ear and Hearing*, 29, 957–970.

McKay, C. M., McDermott, H. J., & Clark, G. M. (1994). Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees. *Journal of the Acoustical Society of America*, 96, 2664–2673.

Miller, J.D. (1961). Word tone recognition in Vietnamese whispered speech. *Word*, 17, 11–15.

Nakata, T., Trehub, S. E., Mitani, C., Kanda, Y., Shibasaki, A., & Schellenberg, E. G. (2005). Music recognition by Japanese children with cochlear implants. *Journal of Physiological Anthropology & Applied Human Sciences*, 24, 29–32.

Nie, K. B., Stickney, G. S., & Zeng, F.-G. (2005). Encoding frequency modulation to improve cochlear implant performance in noise. *IEEE Transactions on Biomedical Engineering*, 52, 64–73.

Nimmons, G. L., Kang, R. S., Drennan, W. R., Longnion, J., Ruffin, C., Worman, T., Yueh, B., & Rubinstein, J. T. (2008). Clinical assessment of music perception in cochlear implant listeners. *Otology & Neurotology*, 29, 149–155.

Peng, S. C., Tomblin, J. B., Cheung, C., Lin, Y.-S., & Wang, L. (2004). Perception and production of Mandarin tones in prelingually deaf children with cochlear implants. *Ear and Hearing*, 25, 251–264.

Peng, S. C., Tomblin, J. B., Spencer, L. J., & Hurtig, R. R. (2007). Imitative production of rising speech intonation in pediatric cochlear implant recipients. *Journal of Speech, Language, and Hearing Research*, 50, 1210–1227.

Pfingst, B. E., Franck, K. H., Xu, L., Bauer, E. M., & Zwolan, T. A. (2001). Effects of electrode configuration and place of stimulation on speech perception with cochlear prostheses. *Journal of the Association for Research in Otolaryngology*, 2, 87–103.

Pretorius, L. L., & Hanekom, J. J. (2008). Free field frequency discrimination abilities of cochlear implant users. *Hearing Research*, 244, 77–84.

Riss, D., Arnoldner, C., Reiss, S., Baumgartner, W. D., & Hamzavi, J. S. (2009). 1-year results using the Opus speech processor with the fine structure speech coding strategy. *Acta Oto-Laryngologica*, 129, 988–991.

Schatzer, R., Krenmayr, A., Au, D. K. K., Kals, M., & Zierhofer, C. (2010). Temporal fine structure in cochlear implants: preliminary speech perception results in Cantonese-speaking implant users. *Acta Oto-Laryngologica*, 130, 1031–1039.

Schouten, J. F., Ritsma, R. J., & Cardoz, B. L. (1962). Pitch of the residue. *Journal of the Acoustical Society of America*, 34, 1418–1424.

Shannon, R. V., Galvin, J. J., 3 rd, & Baskent, D. (2002). Holes in hearing. *Journal of the Association for Research in Otolaryngology*, 3, 185–199.

Smith Z. M., Delgutte B., & Oxenham A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416, 87–90.

Sucher, C. M. & McDermott, H. J. (2007). Pitch ranking of complex tones by normally hearing subjects and cochlear implant users. *Hearing Research*, 230, 80–87.

Tse, J. K., (1978). Tone acquisition in Cantonese: a longitudinal case study. *Journal of Child Language*, 5, 191–204.

Vandali, A. E., Sucher, C., Tsang, D. J., McKay C. M., Chew J. W. D., & McDermott, H. J. (2005). Pitch ranking ability of cochlear implant recipients: a comparison of sound-processing strategies. *Journal of the Acoustical Society of America*, 117, 3126–3138.

Wang, S., Xu, L., & Mannell, R. (2010). Lexical tone recognition in sensorineurally hearing impaired listeners using temporal cues. Paper presented at the American Auditory Society Annual Meeting, Scottsdale, AZ.

Wang, W., Zhou, N., & Xu, L. (2010). Musical pitch and lexical tone recognition with cochlear implants. *International Journal of Audiology*, 50, 270–278.

Wei, C., Cao, K., & Zeng, F. G. (2004). Mandarin tone recognition in cochlear-implant subjects. *Hearing Research*, 197, 87–95.

Wei, W. I., Wong, R., Hui, Y., Au, D. K. K., Wong, B. Y. K., Ho, W. K., Tsang, A., Kung P., & Chung, E. (2000). Chinese tonal language rehabilitation following cochlear implantation in children. *Acta Oto-Laryngologica*, 120, 218–221.

Wei, C., Cao, K., Jin, X., Chen, X., & Zeng, F. G. (2007). Psychophysical performance and Mandarin tone recognition in noise by cochlear implant users. *Ear and Hearing*, 28(2), 62 S–65 S.

Whalen, D. H., & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, 49, 25–47.

Wong, A. O. C., & Wong, L. L. N. (2004). Tone recognition of Cantonese-speaking prelingually hearing-impaired children with cochlear implants. *Otolaryngology-Head and Neck Surgery*, 130, 751–758.

Wong, L. L. N., & Soli, S. D. (2005). Development of the Cantonese HINT. *Ear and Hearing*, 26, 276–289.

Wong, L. L. N., Vandali, A. E., Ciocca, V., Luk, B., Ip, V. W. K., Murray, B., Yu, H. C., & Chung, I. (2008). New cochlear implant coding strategy for tonal language speakers. *International Journal of Audiology*, 47, 337–347.

Wong, P., Schwartz, R. G., & Jenkins, J. J (2005). Perception and production of lexical tones by 3-year-old, Mandarin-speaking children. *Journal of Speech, Language, and Hearing Research*, 48, 1065–1079.

Xu, L., & Pfingst, B. E., (2003). Relative importance of temporal envelope and fine structure in lexical-tone recognition. *Journal of the Acoustical Society of America*, 114, 3024–3027.

Xu, L. & Pfingst, B. E. (2008). Spectral and temporal cues for speech recognition: Implications for auditory prostheses. *Hearing Research*, 242, 132–140.

Xu, L., Tsai, Y., & Pfingst. B. E. (2002). Features of stimulation affecting tonal-speech perception: implications for cochlear prostheses. *Journal of the Acoustical Society of America*, 112, 247–258.

Xu, L., Li, Y., Hao, J. P., Chen, X. W., Xue, S. A., et al. (2004). Tone production in Mandarin-speaking children with cochlear implants: a preliminary study. *Acta Oto-Laryngologica*, 124, 363–367.

Xu, L., Zhang, W., Zhou, N., Lee, C.-Y., Li, Y., et al. (2006). Mandarin Chinese tone recognition with an artificial neural network. *Journal of Otology*, 1, 30–34.

Xu, L., Chen, X., Zhou, N., Li, Y., Zhao, X., et al. (2007) Recognition of lexical tone production of children with an artificial neural network. *Acta Oto-Laryngologica*, 127, 365–369.

Xu, L., Zhou, N., Chen, X., Li, Y., Schultz, H. M., et al. (2009a). Vocal singing by prelingually-deafened children with cochlear implants. *Hearing Research*, 255, 129–134.

Xu, L., Zhou, N., Huang, J., Chen, X., Li, Y., et al. (2009b). Lexical tone development in prelingually-deafened children with cochlear implants. Paper presented at The 12th International Symposium on Cochlear Implants in Children, Seattle, WA.

Xu, L., Chen, X., Lu, H., Zhou, N., Wang S., et al. (2010) Tone recognition and production in pediatric cochlear implants users. *Acta Oto-Laryngologica*, 131, 395–398.

Yuan, M., Lee, T., Yuen, K. C. P., Soli, S. D., van Hasselt, C. A., et al. (2009). Cantonese tone recognition with enhanced temporal periodicity cues. *Journal of the Acoustical Society of America*, 126, 327–337.

Yuen, K. C. P., Yuan, M., Lee, T., Soli, S., Tong, M. C. F., & van Hasselt, C. A. (2007) Frequency-specific temporal envelope and periodicity components for lexical tone identification in Cantonese. *Ear and Hearing*, 28, 107 S–113S.

Zhou, N., & Xu, L. (2008a). Development and evaluation of methods for assessing tone production skills in Mandarin-speaking children with cochlear implants. *Journal of the Acoustical Society of America*, 123, 1653–1664.

Zhou, N., & Xu, L. (2008b). Lexical tone recognition with spectrally mismatched envelopes. *Hearing Research*, 46, 36–43

Zhou, N., Zhang, W., Lee, C.-Y., & Xu, L. (2008). Lexical tone recognition by an artificial neural network. *Ear and Hearing*, 29, 326–335.

Zhou, N., Xu, L. & Lee, C.-Y. (2010). The effects of frequency-place mismatch on consonant confusion. *Journal of the Acoustical Society of America*, 128, 401–409.

Zhu, H., & Dodd, B. (2000). The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language*, 27, 3–42.